ULTRASONOGRAPHIC MEASUREMENT OF CRICOTHYROID SPACE IN SPEECH

Vilkman, Erkki[1]; Takalo, Raija[2]; Määttä, Taisto[3]; Laukkanen, Anne-Maria[4]; Nummenranta, Jaana[1], Lipponen, Tero[1]
From the Departments of [1]Phoniatrics, [2]Roentgenology and [3]Logopedics and Phonetics, University of Oulu, and [4]Department of Communication Research, University of Tampere, Finland;
Corresponding author: Erkki Vilkman, Dept. of Phoniatrics, University of Oulu, FI-90220 Oulu, Finland;
tel. +358 8 3153434, fax. +358 8 3155557, E-mail: Erkki.Vilkman@oulu.fi

ABSTRACT

The physiological background of sentence declination (fundamental frequency, F0, drop) was studied using ultrasound (US) examination of the cricothyroid (CT) space. The US probe was placed anteriorly in the region of middle cricothyroid ligament. The echoes caused by the antero-inferior edge of thyroid and antero-superior edge of the cricoid cartilages were used as points of measurement. The test utterances consisted of three- and five-word sentences. F0, sound level and CT space were measured from recordings. F0 declination and CT space widening showed a phase relationship. E.g., in a long sentence in which the F0 declined from 194 Hz to 85 Hz the CT space change was 4 mm (from 0.83 cm to 1.25 cm). The correlation between the F0 declination and CT space was r=-0.85. The main pitch regulating system connected with CT joint movements seems to contribute to sentence declination production. These biomechanical events can be monitored using the US method.

1. INTRODUCTION

The fundamental frequency (F0) drop (declination) during sentence production is a well known phenomenon. Obviously it is present in all languages. As to the physiological background of this inherent feature of human speech, the significance of subglottal pressure is generally emphasized because subglottal pressure has also been observed to drop during the production of a sentence [1]. From the physiological point of view subglottal pressure is known to affect the pitch to some extent [2].

In general, F0 is considered to be controlled by the cricothyroid (CT) system consisting of the CT joint and the CT muscle. The movements at the CT joint directly influence the length and tension of the vocal folds. This apparatus apparently evolved during evolution only to serve phonatory purposes, i.e. pitch variation [3]. However, the CT system is directly affected by many other mechanisms, such as the thyroarytenoid muscle, external laryngeal muscles and the pull exerted by the trachea on the cricoid cartilage (pitch lowering) [see 4, for a review].

CT muscle function has been investigated electromyographically (EMG) in many contexts. In speech it has been observed to show a relationship with sentence stress and vowel intrinsic F0 production, for instance [5,6]. As to the F0 declination the EMG methods are not informative because after the initial burst of activity the CT muscle shows relatively low level of activity in the absence of extra stress [5]. In addition, the usefulness of EMG methods in voice physiologic studies are severely restricted by the fact that only electrical activity is registered; the problem is that this activity may or may not be connected with relevant biomechanical events.

The outcome of the mutual movements of the cricoid and thyroid cartilage could be observed with direct visualization, e.g., direct fiberoptics which is, however, invasive and subject to measurement problems. Videofluorography would be a method of choice but, unfortunately, the cartilages in guestion are poorly visible which leads to measurement problems, also the radiation risks have to be born in mind.

Measuring the anterior CT distance gives information of the degree of rotation in the CT joint. The CT space is actively diminished by CT muscle activity and its width depends on the balance between the contraction force of thsi muscle and several other forces [4]. Ultrasound (US) examination is favourable because it has no radiation risk, cartilage edges should be readily visible as echoes. The method is widely used for registering dynamic events of the human body. In the present study the method was used to elucidate the physiological background of sentence intonation, especially declination.

2. MATERIAL AND METHODS

Three healthy male subjects were studied at the Department of Roentgenology at Oulu University Hospital. The results for one subject were chosen to be presented in this paper.

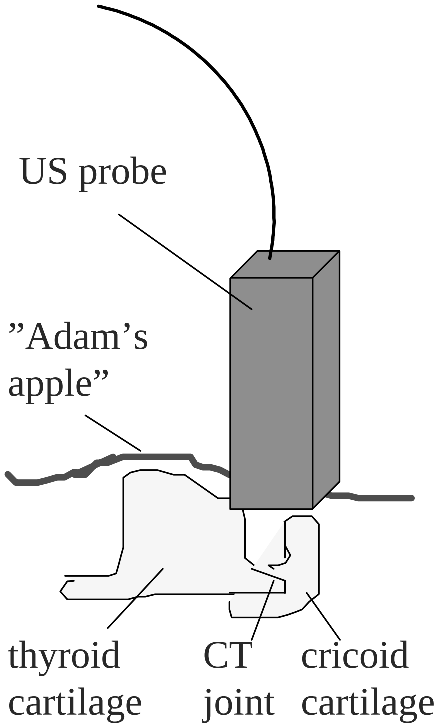

Figure 1. Schematic view of the laryngeal anatomy and the measurement principle. US probe=ultrasound transducer; cranium to the left.

The subject produced two types of sentences: 1) [nal:e meni ma:l:e] (”Teddy went to the country”), 2) [na:le meni ma:l:e laulama:n ma:m:elaulua] (”Teddy went to the country to sing the national anthem”).

During the examination the subject was in supine position. The US examination was performed by means of a real-time scanner (Aloka SSD-2000, Multiview) using a 7.5 MHz linear transducer. The US transducer was placed tightly against the skin anteriorly on the larynx in the midline. The anatomical landmarks and the principle of the examination is shown in Figure 1. A silicone stand-off pad was used, when necessary, to improve the contact.

The US images and the audiosignal were stored on videotape. The frame interval of the videorecorder was 40 ms. The CT space was measured frame by frame (40 ms interval, long sentence) or every second frame (80 ms interval, short sentence. The measurements were performed on monitor screen either manually or using computer cursors. The values were calibrated against recorded scales. The F0 and sound level on a relative scale (SL) analyses were carried out using a microcomputer-based analysis package (CSL 4300) with 20 ms windows.

3. RESULTS AND DISCUSSION

Many factors affect the accuracy of the measurement. Firstly, the landmarks, i.e. the echoes, used in the measurements may be affected by soft tissue displacement. Secondly, the contact between the skin and the probe has to be very tight. Using a stand-off pad improved the quality of the images in complicated cases. Thirdly, the limited frame rate of the video recorder may cause some blurring of the images during the fast changes. As to measurement technique both the manual measurement and the measurement with the computer cursors have strengths and weaknesses. The correlation taken from a sample measured with the two methods was r=0.85 (n=16). Reasonably reliable results can, thus, be obtained with both methods. The advantage of measuring directly on the video monitor screen is that the investigator has the opportunity to see the dynamics of the movements.

Examples of the results are presented in Figures 2 and 3. The timing of the measurements are hampered by the different framing of the video and audiosignals. However, it is interesting to note that the initial F0 rise of the utterances showed some deviation from the corresponding CT distance in terms of timing. The F0 already dropped while the CT distance still remained at the minimum level. In Figure 3 the F0 rise took place well before the CT distance reached its minimum level which might be explained by the relatively high F0 implying a F0 raising technique related to register control and/or the anteroposterior gliding in the CT joint also causing increased longitudinal tension of the vocal folds [4].

All in all, the minimum level of the CT space had a longer duration than the F0 peak. This finding may reflect the interplay between the laryngeal muscles: the CT and the thyroarytenoid muscles are anatomically antagonistic but functionally agonistic in terms of F0 control [c.f. 2]. The timing discrepancy might be related to this biomechanical arrangement. On the other hand, the CT muscle consists of two separate parts which cause mainly rotation (pars recta) or anteroposterior gliding (pars obliqua) at the CT joint [4]. These findings naturally call for further studies, however, at the moment it seems that US imaging of the CT space might be helpful in understanding the complex interaction of the laryngeal muscles in pitch regulation at the biomechanical level.

The F0 declination can be seen in Figures 2 and 3. In both cases the CT space gradually opens well in relationship with the dropping F0. This is most likely a consequence of slowly relaxing CT muscle contraction. This assumption is based on the fact that the true pitch lowering mechanism, namely, the tracheal pull is always present and would cause a rapid F0 drop to a

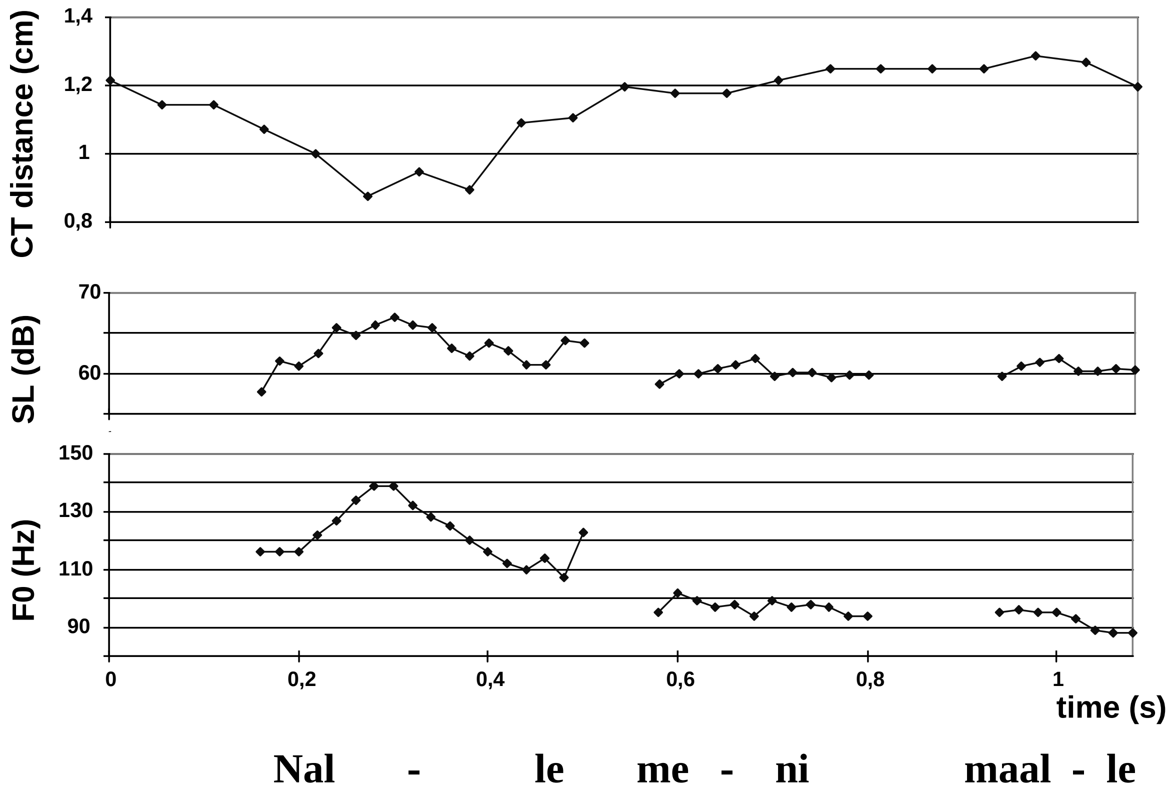

Figure 2. Cricothyroid (CT) distance, sound level (SL) and fundamental frequency (F0) for the sentence [nal:e meni ma:l:e]. The first word stressed.

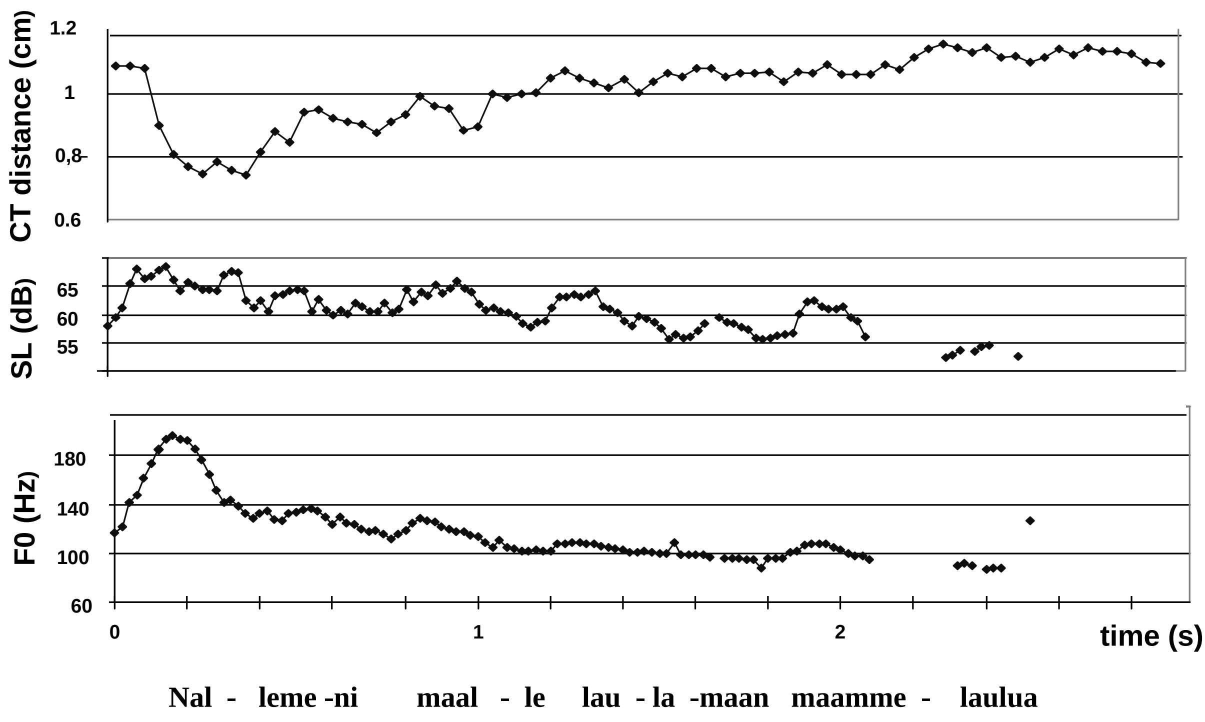

Figure 3. Cricothyroid (CT) distance, sound level (SL) and fundamental frequency (F0) for the sentence [nal:e meni ma:l:e laulama:n ma:m:elaulua].

very low level in case the CT muscle would be totally relaxed [see 4, for a review].

In the short sentence (Fig. 2) the F0 variation from the peak to the value measured at the end of the utterance (about 50 Hz) was related to a CT space change of 0.4 cm., i.e. 13 Hz/mm. In the long sentence the corresponding CT space change was also 0.4 cm but the F0 change larger (about 100 Hz, 25 Hz/mm). This reflects the variability of laryngeal settings in F0 control. In the long sentence (Fig. 3) the overall correlation between F0 and CT distance was r=-0.48 (p=0.000). When only the slowly dropping part of the F0 declination curve was taken into account the correlation between CT space and F0 was r=-0.85.

CONCLUSIONS

The US method seems to open new views for further studies of the mutual movements of the cricoid and thyroid cartilages.

The finding that the CT space increases gradually with F0 declination implies the contribution of the CT system to this phenomenon. A reminder seems necessary here as to that the physiology of declination should not be undersood as a dichotomy, that is to say that the question is not whether either subglottal pressure or CT system is the crucial factor in its control. Both systems are present and are basically acting to the same direction, besides it obvious that the significance of these factors varies individually and also from one situation to another.

4. REFERENCES

[1] Cruttenden A. Intonation. Cambridge textbooks in linguistics. Cambridge University Press, Cambridge 1986.

[2] Titze, I. Principles of voice production. Prentice-Hall, Englewood Cliffs 1994.

[3] Wind, J. Phylogeny of the human vocal tract. In: Harnad, S., Steklis, H., Lancaster, J. eds. Origins and evolution of language and speech. Ann NY Acad Sci 1976; 280:612-620.

[4] Vilkman, E., Sonninen, A., Hurme, P., Körkkö, P. External laryngeal frame function in voice production revisited: a review. J Voice 1996; 10:78-92.

[5] Vilkman, E., Aaltonen, O., Raimo. I., Ignatius, J., Komi, P: On stress production in whispered Finnish. J Phonetics 1987; 15:157-168.

[6] Vilkman, E., Aaltonen, O., Raimo, I., Arajärvi, P., Oksanen, H. Intrinsic pitch of vowels - a complicated problem with an obvious solution. pp. 159-166. In: Gauffin, J., Hammarberg, B. eds. Vocal fold physiology. Singular, San Diego 1991.